

J Autism Dev Disord (2011) 41:1256–1266
DOI 10.1007/s10803-010-1157-x

ORIGINAL PAPER

Diagnosing Autism Spectrum Disorders in Adults: the Use of Autism Diagnostic Observation Schedule (ADOS) Module 4

Jojanneke A. Bastiaansen · Harma Meffert · Simone Hein · Petra Huizinga ·
Cees Ketelaars · Marieke Pijnenborg · Arnold Bartels · Ruud Minderaa ·
Christian Keyzers · Annelies de Bildt

Published online: 14 December 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Autism Diagnostic Observation Schedule (ADOS) module 4 was investigated in an independent sample of high-functioning adult males with an autism spectrum disorder (ASD) compared to three specific diagnostic groups: schizophrenia, psychopathy, and typical development. ADOS module 4 proves to be a reliable instrument with good predictive value. It can adequately discriminate ASD from psychopathy and typical

development, but is less specific with respect to schizophrenia due to behavioral overlap between autistic and negative symptoms. However, these groups differ on some core items and explorative analyses indicate that a revision of the algorithm in line with Gotham et al. (J Autism Dev Disord 37: 613–627, 2007) could be beneficial for discriminating ASD from schizophrenia.

Keywords Autism · Adults · ADOS · Diagnosis · Schizophrenia · Psychopathy

J. A. Bastiaansen (✉) · H. Meffert · C. Keyzers
Social Brain Lab, Department of Neuroscience, University
Medical Center, Groningen, Hanzeplein 1, 9700 RB Groningen,
the Netherlands
e-mail: j.bastiaansen@med.umcg.nl

J. A. Bastiaansen · S. Hein · P. Huizinga
Autism Team North Netherlands, Lentis, Laan Corpus den
Hoorn 102-2, 9728 JR Groningen, the Netherlands

H. Meffert · A. Bartels
FPC Dr. S. van Mesdag, Forint, Helperlinie 2, 9722 AZ
Groningen, the Netherlands

C. Ketelaars
Jeugd en Autisme, Dimence, Burgemeester Roelenweg 9, 8021
EV Zwolle, the Netherlands

R. Minderaa · A. de Bildt
Department of Psychiatry, University Medical Center
Groningen, University of Groningen, Hanzeplein 1, 9700 RB
Groningen, the Netherlands

M. Pijnenborg
Psychosociaalcluster, GGZ Drenthe, Dennenweg 9, 9404 LA Assen,
the Netherlands

C. Keyzers
Social Brain Lab, Netherlands Institute for Neuroscience, Royal
Netherlands Academy of Arts and Sciences (KNAW),
Meibergdreef 47, 1105 BA Amsterdam, the Netherlands

Introduction

Although for a diagnosis of an autism spectrum disorder (ASD) symptoms should be present from infancy or early childhood, the disorder may not be detected until later because of several reasons: a well-structured support system, compensation for limitations through high intelligence, the presence of more subtle autistic symptoms, and confusion with or overshadowing by another psychiatric disorder (Kan et al. 2008; Wing and Potter 2002, see also www.dsm5.org). Partly due to increasing knowledge of milder forms of autism and more awareness that autistic conditions can be found in individuals of high ability, ASDs are starting to become more widely recognized in adults (Brugha et al. 2009; Fombonne 2005; Kan et al. 2008; Wing and Potter 2002). In clinical practice, we notice a growing demand for diagnostic procedures concerning ASD in adults. However, there is no established diagnostic tradition for ASD in older individuals. It is very challenging to disentangle social and communicative problems associated with ASD from the often complicated clinical picture in adulthood, especially when developmental information is

unavailable. Standardized instruments are needed that can facilitate the diagnostic process. Poor self-referential cognition present in many individuals with ASD may hamper self-report measures of autistic symptoms (Johnson et al. 2009; Lombardo et al. 2007). Therefore, observation of the individual during social interaction is important in addition to information about difficulties experienced in daily life.

The Autism Diagnostic Observation Schedule (ADOS, Lord et al. 2000) is a standardized instrument that assesses social interaction, communication, and imagination during a semi-structured interaction with an examiner. The ADOS includes four modules suited for individuals with different developmental and language levels, ranging from children with no expressive language to older and verbally more capable individuals. The psychometric properties of modules 1–3 are well-studied and present the ADOS as a reliable and valid instrument to assess the presence of ASD in children (de Bildt et al. 2004; Gray et al. 2008; Lord et al. 2000; Noterdaeme et al. 2000; Papanikolaou et al. 2009). Module 4 was developed for adolescents and adults with fluent speech. In the original paper on the ADOS, Lord et al. (2000) included module 4 administrations for adolescents and young adults with autism (AD, $n = 16$), PDD-NOS ($n = 16$) and with various other diagnoses ($n = 15$). Their results indicate that, after training as described in the manual (Lord et al. 1999), ADOS module 4 can be used effectively to distinguish between autism spectrum and non-spectrum, and to a lesser degree between AD and PDD-NOS. Thus far, no further specific studies into the value of module 4 have been reported. When establishing a diagnosis, clinicians need to rule out specific conditions that can cause similar symptoms. Because the control group in Lord's study (2000) was relatively small and very diverse with respect to diagnosis, it is still unclear to what extent ADOS module 4 can support such differential diagnostics.

One disorder that shares symptoms with autism is schizophrenia. Kanner (1943) even borrowed the term autism from Eugen Bleuler, who used it to describe withdrawal from contact with the outside world in adults with schizophrenia (1911). Although autism and schizophrenia have different developmental trajectories, cross-sectionally their clinical presentations overlap (Frith and Happé 2005; Goldstein et al. 2002; Volkmar and Cohen 1991). Especially individuals with schizophrenia and negative symptoms show many of the same social deficits as adults with autism (Frith and Happé 2005). Autism also shares features with psychopathy, a personality disorder which partly overlaps with antisocial personality disorder (APD). Besides poor behavioral control and a disregard for the rights of other people, individuals with

psychopathy have deficits in the emotional and interpersonal domain, such as insensitivity or lack of empathy towards other people. Impairments in empathy are also central to ASD, characterized by a cognitive impairment to take the perspective of other people (Baron-Cohen and Wheelwright 2004; Gillberg 1992). Rogers et al. (2006) indicate that there could be a subgroup of people with ASD that have additional callous-unemotional traits reminiscent of psychopathy. Others report that some individuals with ASD may *seem* cold and heartless, because they are unaware of how their behavior affects other people, which could lead to a diagnosis of APD or psychopathy by mistake (Bartels and Bruinsma 2008; Howlin 2000; Kohn et al. 1998). Especially in forensic settings, it is important to differentiate ASD from psychopathy, because they require different approaches. It should be noted, however, that unlike in psychopathy there is little evidence of any excess of crimes among people with autism (Howlin 2000).

The current study will examine the psychometric properties of ADOS module 4 by including relatively homogeneous non-autistic groups: a group of adult males with schizophrenia and marked negative symptoms, males with psychopathy, and typically developing males. Analyses will center on the original ADOS algorithm (Lord et al. 2000), based on the operationalization of the DSM-IV and ICD-10 criteria for autistic disorder (American Psychiatric Association 1994; World Health Organization 1993), but will also include some preliminary analyses based on revised algorithms for the ADOS. In line with proposals for the revision of the DSM (www.dsm5.org), the revised algorithms of the ADOS for modules 1–3 synthesize the items from the original social interaction and communication domains into the new domain Social Affect (SA, Gotham et al. 2007). This new notion of communicative and social behaviors as a single set of symptoms is supported by recent studies showing that non-verbal communication and social items load onto the same factor (Constantino et al. 2004; Lord et al. 1999; Robertson et al. 1999; van Lang et al. 2006). In addition, the revised algorithms include restricted and repetitive behaviors (RRB) as opposed to the original algorithm. Although the narrow time frame of the ADOS might not provide adequate opportunity to measure these behaviors (Lord et al. 2000), they seem to make an independent contribution to diagnostic stability (de Bildt et al. 2009; Lord et al. 2006). While adults with ASD may have a slightly different behavioral phenotype compared to children (Gotham et al. 2007), the core difficulties persist in adulthood (Seltzer et al. 2004; Shattuck et al. 2007). Therefore, it is of interest to explore the utility of this promising new metric in our adult population.

Methods

Participants

Thirty-two adult males with an ASD were recruited via local mental health organizations (mainly through the specialized Autism Team North Netherlands of Lentis, Groningen, the Netherlands), and through mailing lists for high-functioning individuals with ASD. Six individuals with ASD were recruited from a local forensic clinic (FPC Dr. S. van Mesdag, Groningen, the Netherlands). The participants were considered to be high-functioning by their clinicians and none had an IQ score below 70. All participants were diagnosed with an ASD by a clinical psychologist or psychiatrist according to DSM-IV-TR criteria ($n = 8$ AD, $n = 17$ AS, $n = 13$ PDD-NOS), based on review of developmental history, current daily functioning, and observation. For this study, the ASD group will be investigated as one diagnostic entity along a continuous dimension of severity for two reasons. First, it is proposed for the near future that distinctions will no longer be made among different types of autism in clinical practice, because they have proven to be “inconsistent over time and place, and to be associated more with severity, language level, and intelligence than specific features” (www.dsm5.org). Individuals with autism and PDD-NOS have also shown qualitatively similar behavioral patterns on the ADOS with varying degrees of severity (Lord et al. 2000). Second, investigating the subtypes would lead to overly small subgroups.

Eighteen adult males with schizophrenia and predominantly negative symptomatology, mainly outpatients, were selected by a specialized local mental health organization (Psychosencluster, GGZ Drenthe, Assen, the Netherlands). Diagnosis was confirmed by a structured clinical interview, the Dutch version of the Schedules of Clinical Assessment in Neuropsychiatry developed by the WHO (SCAN 2.1, Giel and Nienhuis 1996). Current symptomatology was assessed by the Positive and Negative Syndrome Scale (PANNS, Kay et al. 1987).

The psychopathy group consisted of 16 males recruited from two forensic psychiatric clinics (FPC Dr. S. van Mesdag and FPC Veldzicht). As part of the standard clinical procedure, these individuals were assessed with the Psychopathy Checklist Revised (PCL-R), an instrument widely used for the diagnosis of psychopathy (e.g. Hare 1991). Two diagnosticians obtained consensus on this instrument after separately scoring the items using file information extended with, if necessary, a semi-structured interview.

The typically developing group consisted of 21 typically developing males, who were interviewed to verify that first-degree relatives did not have an ASD or a history of

Table 1 Group characteristics

	<i>N</i>	Mean	Stdev	Range
ASD				
Age	38	31.82	11.24	18–66
IQ	29	101.14	14.67	73–133
Schizophrenia				
Age	18	37.00	10.73	19–61
IQ	18	89.17	13.89	68–112
Psychopathy				
Age	16	39.00	10.67	23–60
IQ	15	92.73	16.10	63–117
Controls				
Age	21	34.24	9.14	21–53
IQ	21	97.19	16.37	73–128

IQ scores were based on the Groninger Intelligence Test 2 (GIT2, Luteijn and Barelds 2004), except for four individuals with ASD who were administered the Wechsler Adult Intelligence Scale (WAIS, Wechsler 1997) and nine individuals with ASD for whom IQ scores were not obtained (they only took part in the ADOS part of the research project). For these cases, IQ was estimated to be in the normal range based on former IQ tests and clinical impression/daily functioning. GIT 2 scores for one individual with psychopathy were deemed unreliable and discarded

psychosis. Age and IQ was matched with the participants with ASD who also took part in the neuroimaging part of the study ($n = 21$). There are no significant differences between the groups in terms of age and IQ. For an overview of the group characteristics see Table 1.

Measures and Procedure

Administration of the ADOS was part of the standard procedure of two large neuroimaging studies into the neural basis of empathy conducted in the Social Brain Laboratory (www.bcn-nic.nl/socialbrain.html). All participants gave written informed consent. The studies were approved by the Institutional Review Board of the University Medical Center Groningen (METc). The administration of ADOS module 4 included all standard activities and the optional daily living items to obtain relevant background information. The interviews were administered and scored by trained and certified psychologists. In total, five raters participated in the project including two certified ADOS trainers (AdB, SH). To ensure that agreement between raters remained at the high level requested by the ADOS, we discussed (fragments of) videotapes in two-monthly group meetings. The interviews were scored for consensus from videotape in changing pairs of raters, but included the examiner in the far majority of cases. In contrast to the second rater, the examiner was not blind to clinical diagnosis. The consensus scores were determined on the basis of the video-recording through a discussion in

which the judgment of each rater was weighted equally. We only made an exception to this procedure when there was major disagreement (0 vs. 2) for the items B1 (Eye Contact) and B2 (Facial Expressions). Then, we gave priority to the examiner's opinion, because we anticipated that these items might be difficult to judge from videotape alone. Fortunately, due to the high quality of the video-recordings, there was major disagreement in only two out of 93 administrations for eye contact, while for facial expressions such disagreement never occurred. Therefore, it is unlikely that the examiner's previous knowledge influenced the consensus scores.

Design and Analysis

Algorithms

In this paper, we will use the terms “original algorithm” when referring to the standard algorithm (Lord et al. 2000) and “revised algorithm” when referring to the application of the revised algorithm based on Gotham et al. (2007). To reach a classification of AD or ASD on the original algorithm of the ADOS, an individual needs to meet thresholds for the communication domain (COM), the social interaction domain (SOC), and for the summation of these two domains (COMSOC), but not for the restricted and repetitive behaviors domain (RRB, Lord et al. 1999). For the revised algorithm, classification is based on solely thresholding the SARRB domain, which combines social, communication, and restricted behavior items. Since algorithm items across modules 3 and 4 are comparable and our sample size does not permit independent factor analyses in order to establish specific algorithm items, we applied the revised algorithm for module 3 to our group of high-functioning adults to calculate domain scores and a total score. In line with the explanation on the original algorithm, scores of 3 were converted to 2, and all scores other than 0–3 were treated as 0.

Interrater Agreement

Interrater agreement was assessed on the original algorithm at the level of ADOS classification, domains, and items. Agreement between raters at the level of diagnostic classification (AD, ASD, nonspectrum) was calculated through Cohen's weighted kappa in addition to the percentage of agreement. Cohen's kappa takes into account the agreement that can occur by chance between two raters and is therefore more stringent than the mere calculation of the percentage of times the raters' scores lead to the same ADOS classification. Interrater agreement on the domains and the total score was calculated by means of intraclass correlations (ICC). ICC scores represent correlations across

pairs of raters and are higher the more consistent the scores across two different raters are. ICC scores, internal consistency and correlations could not be reliably calculated for the RRB domain, because variance was too limited: for four out of the five items less than five subjects scored different from zero. To assess interrater agreement for separate items, we used mean linearly weighted Cohen's kappa's in line with Lord et al. (2000). Cohen's linearly weighted kappa takes into account the agreement between two raters occurring by chance and considers the difference between a score of zero and one to be smaller than a difference between zero and two. Item B3 was ignored because its score depends on items A9, B1 and B2. In addition, only items were included for which more than five subjects scored different from zero (excluding nine items: A1, A3, A5, D1, D2, D3, D5, E1, E2).

Internal Consistency

To measure the internal consistency of the original and the revised domains, we applied Cronbach's alpha. This statistic increases as the intercorrelations among test items within a domain increase.

Comparison of Domain Means

We used an ANOVA for each scale of both algorithms with fixed factor group, followed up by Tukey's HSD post-hoc comparisons. We performed one-tailed Mann–Whitney tests to examine whether the forensic ASD group scored higher than the psychopathy group. To compare group differences at item level, we performed a MANOVA with fixed factor group on all items except the previously mentioned nine items that had limited variance and item B3. Post-hoc tests were performed for those items that showed a significant group effect.

Criterion-Related Validity

Here, criterion-related validity refers to the degree to which the outcome on the ADOS instrument is in agreement with the clinical diagnosis of having ASD or not. We used logistic regression to measure the success of both algorithms in predicting whether a participant received a diagnosis of ASD in clinical practice. Because ADOS classification is based on COM and SOC for the original algorithm and on the combined SARRB domain for the revised algorithm, we used these domains as predictors in two separate analyses. Logistic regression provides information on the sensitivity and specificity for the fixed cut-off point used in clinical practice. Receiver Operating Characteristic (ROC) curves provide information on the sensitivity and specificity of all other possible scores. In

addition, it provides an Area under the Curve statistic (AuC), which represents the overall level of agreement between criterion (i.e. clinical diagnosis of ASD) and instrument (i.e. ADOS). The higher the AuC, the higher the probability that a randomly chosen participant with ASD will have a higher score on the instrument than a randomly chosen participant without ASD.

Correlations with Participant Characteristics

To investigate the relationship of domain scores with participant characteristics, we calculated bivariate correlations for the patient groups between domain scores, and age, IQ, and scores on the negative scale of the PANNS (schizophrenia only).

Results

Interrater Agreement

Interrater agreement at the level of ADOS classification was 81.7% with Cohen's adjusted weighted kappa 0.66, which corresponds to good or substantial agreement (Landis and Koch 1977). When merging the ADOS-classifications AD and ASD (based on the proposed criteria for DSM V) the agreement increased to 89.2% with kappa 0.73. Intraclass correlations (ICC, Table 2) show high interrater agreement on SOC and COMSOC, and good agreement on COM. Mean agreement across the items was 81.7% with mean weighted kappa 0.66. Weighted kappa's exceeded 0.60 for 14 out of the 21 items with the remainder exceeding 0.50.

Internal Consistency

For the original algorithm, the internal consistency is high for SOC (Cronbach's $\alpha = 0.84$), but rather low for COM ($\alpha = 0.52$). This indicates that the items of that domain do not intercorrelate well in our population. Item A4 (Stereotyped Language) performed the worst and its deletion from COM increased alpha to an acceptable level ($\alpha = 0.60$). The reorganization of communication and social interaction items in the SA domain of the revised algorithm creates a consistent domain ($\alpha = 0.87$).

Comparison of Domain Means

Original Algorithm

All three domains and the total score showed a significant difference between the groups (Table 3). Tukey post-hoc comparisons show that for COM, SOC, and COMSOC, the ASD group scores significantly higher compared to the psychopathy group and the control group, but not compared to the schizophrenia group. The schizophrenia group scored significantly higher than the control group on COM, and higher than both the psychopathy and the control group on SOC and COMSOC. For RRB, the ASD group scored significantly higher than the control group, while there was a trend compared to the psychopathy group ($p = .06$). The forensic subgroup with ASD ($n = 6$) scored higher than the group with psychopathy on all domains (data not shown).

Revised Algorithm

Both domains and the total score showed a significant difference between the groups (Table 3). Tukey post-hoc comparisons indicated that the ASD group scored significantly higher compared to the psychopathy group and the control group on SA, and there was a trend in comparison to the schizophrenia group ($p = .06$). The schizophrenia group scored significantly higher than the control group. For RRB, the ASD group again scored significantly higher than the psychopathy and control groups, but there was no significant difference with the schizophrenia group. For the total SARRB score, the ASD group scored significantly higher than the psychopathy, the control group, and the schizophrenia group, making it the only score for which the ASD group significantly differs from the schizophrenia group. The forensic subgroup with ASD ($n = 6$) scored higher than the group with psychopathy on all domains (data not shown).

Group Comparison at Item Level

The multivariate test showed that there was a significant main effect of group, $F(66,210) = 1.688$, $p < .005$. Results for the univariate tests are visually presented in Figure 1.

Only four out of 22 items did not differ significantly between the groups. The majority of the remaining items

Table 2 Intraclass correlations for interrater agreement

N	Social interaction	Communication	Social-Communication
93	0.92 (Lord: 0.88–0.97)	0.79 (Lord: 0.74–0.90)	0.92 (Lord: 0.84–0.98)

Interrater agreement as mentioned for Lord et al. (2000) represents the range for all four modules

Table 3 Summary statistics based on the original and revised algorithms

	ASD (<i>n</i> = 38)	Schizophrenia (<i>n</i> = 18)	Psychopathy (<i>n</i> = 16)	Control (<i>n</i> = 21)	<i>F</i> (3,89) ^a	Post-hoc tests ^b
Original algorithm						
Communication domain (COM)						
Mean	2.55	1.94	1.00	0.71	8.82***	ASD > P**
SE	0.27	0.40	0.26	0.22		ASD > TD***
Range	0–7	0–7	0–3	0–3		S > TD*
Social interaction domain (SOC)						
Mean	6.13	4.28	1.50	0.95	18.69***	ASD > P***
SE	5.97	7.62	0.38	0.25		ASD > TD***
Range	0–13	0–13	0–5	0–4		S > P*/TD**
Social-communication total (COMSOC)						
Mean	8.68	6.22	2.50	1.67	19.69***	ASD > P***
SE	0.80	0.90	0.52	0.43		ASD > TD***
Range	0–17	2–17	0–8	0–7		S > P*/TD**
Restricted and repetitive behaviors domain (RRB)						
Mean	0.37	0.33	0.00	0.00	4.07**	ASD > TD*
SE	0.11	0.11	0.00	0.00		
Range	0–2	0–1	0–0	0–0		
Revised algorithm						
Social affect domain (SA)						
Mean	7.39	4.67	1.94	1.48	14.97***	ASD > P***
SE	0.78	0.95	0.36	0.38		ASD > TD***
Range	0–17	0–17	0–5	0–5		S > TD*
Restricted and repetitive behaviors domain (RRB)						
Mean	0.97	0.56	0.25	0.10	7.73***	ASD > P**
SE	0.16	0.15	0.11	0.07		ASD > TD***
Range	0–4	0–2	0–1	0–1		
SARRB total						
Mean	8.37	5.22	2.19	1.57	17.50***	ASD > S*
SE	0.84	0.95	0.40	1.81		ASD > P***
Range	0–17	0–17	0–6	0–5		ASD > TD***

^a For the univariate comparisons (main effects of group), *F* scores and their respective significance levels are reported: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

^b Using the same symbols, Tukey's HSD post-hoc comparisons are reported in the outer right column to indicate what diagnostic groups are significantly different from each other. *ASD* refers to autism spectrum disorder, *P* psychopathy, *TD* typically development, *S* schizophrenia

showed a (almost) significant difference between the ASD group compared to the psychopathy and control groups, but not compared to the schizophrenia group. On some of these items the schizophrenia group also scored significantly higher than the psychopathy and/or control group: B2 (Facial Expressions), B6 (Empathy/Comments on Others' Emotions), and B7 (Insight). Only three items distinguished the ASD from the schizophrenia group: A4 (Stereotyped Language), B10 (Quality of Social Response), and B12 (Overall Quality of Rapport). In addition, there was a trend for the ASD group to score higher than the schizophrenia group on item B11 (Amount of Reciprocal Social Communication, $p = .07$).

Individuals with psychopathy scored comparable to the control group.

Criterion-Related Validity

The ADOS was able to correctly classify 74.2% of the cases in our sample as having ASD or not (based on the clinical diagnosis assigned). Logistic regression analysis showed that SOC ($p < .005$) but not COM ($p = .27$) made a significant contribution in predicting whether a participant in our sample had a clinical diagnosis in the autism spectrum or not (Table 4). The SARRB domain significantly contributed to prediction (Table 4, $p < .005$).

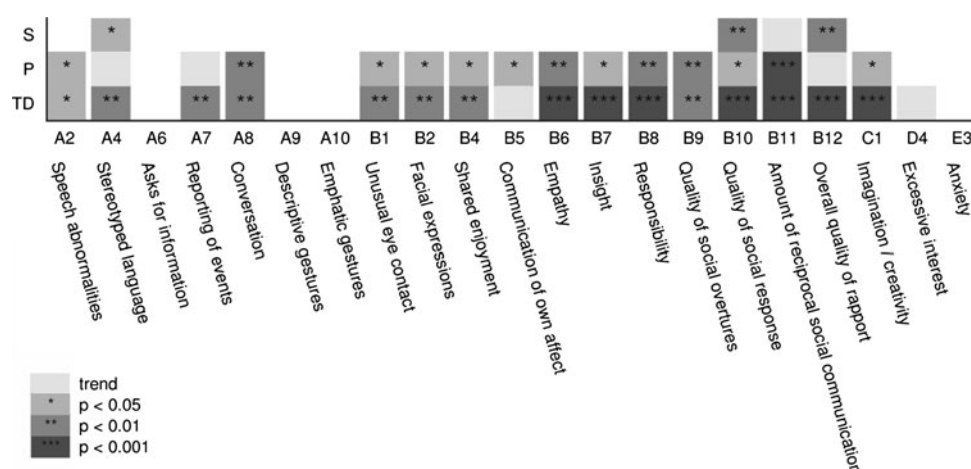


Fig. 1 Between-group Comparisons at Item Level. Post-hoc comparisons of the ASD group versus the other three groups at item level (S schizophrenia, P psychopathy, TD typical development). Dark grey boxes filled with *** represent a statistically significant difference at $p < .001$. Middle grey boxes filled with ** represent a statistically

significant difference at $p < .01$. Light grey boxes filled with * represent a statistically significant difference at $p < .05$. Unfilled light grey boxes represent a statistical trend ($p < .1$). In all these cases, the mean of the ASD group was higher compared to the respective group

Table 4 Logistic regression analyses for criterion-related validity

ADOS		Clinical classification						
Algorithm	Domain	B	SE	Wald	df	Sig.	Odds ratio	95% C.I.
Original	Communication	0.21	0.20	1.21	1	0.271	1.23	0.85–1.79
	Social interaction	0.32	0.09	11.77	1	0.001	1.38	1.15–1.66
Revised	SARRB	−0.29	0.07	18.52	1	0.000	1.33	1.17–1.52

Odds ratios express how much the probability that an individual has received a clinical diagnosis of ASD increases with augmenting scores on the ADOS

The odds ratios presented in Table 4 indicate that augmenting scores of one point on SOC or SARRB, increase the probability that the individual has received a clinical diagnosis of ASD by 38 and 33%, respectively.

ROC curves for the original and revised algorithms resulted in AuC values of .812 and .796, respectively (1 = perfect agreement). This indicates that in general the ADOS scores quite adequately predicted whether someone had a clinical diagnosis of ASD or not. Application of the standard cut-off for autism spectrum on the original algorithm (i.e. 7) gives only moderate sensitivity (0.61) but good specificity (0.82) in our sample. Lowering the threshold to 6 increases the sensitivity (0.68) and keeps the specificity at the same level (0.82). Lowering the threshold to 5 increases the sensitivity further (0.79), but it decreases the specificity (0.73). For the revised algorithm, a cut-off of 5 seems optimal in the current population with adequate sensitivity (0.71) and specificity (0.82).

Correlations with Participant Characteristics

There were no significant correlations between the domain scores, and IQ or age for the groups with ASD,

schizophrenia, nor psychopathy (data not shown). In the group with schizophrenia, the presence of negative symptoms as measured by the PANNS correlated positively with SOC ($r = 0.59$, $p < .05$) but not COM ($r = 0.12$). The PANNS also correlated positively with SA ($r = 0.66$, $p < .005$). Thus, the more negative symptomatology an individual with schizophrenia had, the higher his scores on the ADOS. PANNS scores correlated in particular with items that are similar to negative symptoms, such as (flat) facial expressions (B2, $r = 0.59$, $p < .05$), (lack of) shared enjoyment (B4, $r = 0.81$, $p < .01$), (lack of) asking the examiner for information (A6, $r = 0.66$, $p < .01$), and (difficulty with) communication of own emotions (B5, $r = 0.53$, $p < .05$).

Discussion

Systematic instruments are needed that can facilitate the complicated diagnostic process concerning ASD in adults. The current study is the first that examined the psychometric properties of ADOS module 4 in an independent sample of high-functioning adult males with an established

clinical ASD diagnosis compared to meaningful and relatively homogeneous clinical and non-clinical groups. Our findings show that ADOS module 4 is a reliable instrument. At all levels (i.e. classification, domains and items) raters obtained substantial agreement. In addition, ADOS module 4 has good general criterion-related validity. It is able to correctly classify the majority of individuals and higher scores on the ADOS predict a higher probability of having a clinical ASD diagnosis. The high Areas under the Curve are further indications that ADOS scores can predict whether an individual actually has an ASD. Furthermore, group comparisons between the ASD and other groups show that the ADOS is valuable in differentiating between ASD, and psychopathy and typical development. The distinction between psychopathy and ASD even holds when only taking into account forensic individuals with ASD (although the group size was rather small to perform such an analysis). The finding that ASD and psychopathy are so well-discriminated by means of ADOS scores is promising for forensic psychiatric settings.

Another finding is the similarity between ASD and schizophrenia with respect to ADOS scores. Clearly, individuals with schizophrenia and marked negative symptoms show behavior that is very similar to ASD (Frith and Happé 2005). Some patients with schizophrenia even have autistic-like symptoms that covary with negative symptoms (Sheitman et al. 2004). In line with these data, we show that the degree of negative symptomatology correlates significantly with ADOS scores, in particular with items resembling negative symptoms, such as (lack of) directed facial expressions and shared enjoyment. This resemblance makes it difficult for an observational instrument such as the ADOS to differentiate these groups on that behavior (see Reaven et al. 2008 for a similar finding in children with childhood-onset schizophrenia). The findings underscore previous recommendations of using a comprehensive assessment that incorporates information on daily functioning and early development with direct observation to reach a clinical diagnosis (Lord et al. 1999). Nevertheless, four items did show a difference between these groups: individuals with ASD use more stereotyped language, less reciprocal social communication, and display qualitatively poorer social responses and overall rapport. This suggests that core social items and stereotyped language discriminate individuals with ASD from those with schizophrenia.

Although findings are preliminary, the revised SARRB domain, which combines social, communication and repetitive behavior items, seems promising in this and other respects. It not only discriminates ASD from all other groups including schizophrenia, but also has high internal consistency, and does well in identifying ASDs: a higher score on this domain predicts a higher probability of a

clinical ASD diagnosis with 33% per additional point. Another positive indication for the revised algorithm is the confirmation that stereotyped language fits better with the RRB factor than with the original communication domain. Notwithstanding the caution of interpreting ASDs in adults in exactly the same way as in children, the revised algorithm as developed for modules 1–3 seems promising for module 4 as well. More research is needed in a larger sample containing individuals with more severe autistic symptoms and lower levels of daily functioning to further investigate the revised algorithm.

A marked finding is the limited role of the original communication domain in the identification of ASDs in this sample. Despite group differences between ASD and psychopathy/typical development, the communication domain does not predict a clinical ASD diagnosis. Combined with its low internal consistency, the communication domain as such does not seem to add to the validity of ADOS module 4 in the current sample. However, when communication items are incorporated in the revised algorithm, a consistent scale (SA) emerges that is valuable in the diagnostic procedure for ASD. Similarly, although restricted and repetitive behaviors were rare in our ASD sample, their contribution to SARRB supports the distinction of ASD from schizophrenia. The relatively short duration of the ADOS interview naturally could have played a role in the paucity of RRBs (Lord et al. 1999). However, combining these two findings also fits the general clinical picture: in adolescents and adults with ASD there is a greater prevalence of impairment in non-verbal communication and social reciprocity than in verbal communication or repetitive behaviors and stereotyped interests (Shattuck et al. 2007). In fact, repetitive behaviors decline most strongly with age (Seltzer et al. 2003). Apart from ageing, individuals in our sample might have had relatively more intact verbal skills from the outset as they were all considered to be high-functioning. Stereotyped language, however, does differentiate the ASD group from all other groups in our sample. This may be typical of our high-functioning group, because idiosyncratic language and language complexity are positively associated (Volden and Lord 1991). Cultural differences in the use of gestures might also have played a role. Typically developing adults in our sample, for instance, used few emotional and only occasional descriptive gestures themselves.

The sensitivity in our sample was rather low (0.61), which means that not every individual with a clinical diagnosis of ASD obtained a concurrent classification on the ADOS. It is probable that the characteristics of our group played a role in this. Our sample consisted of high-functioning individuals that signed up for an extensive research project. They are probably situated at the milder end of the spectrum and some might have been able to

(partly) compensate some behavior due to their high intelligence. Resulting relatively low scores make it difficult for the ADOS to identify these individuals. Our findings resemble the outcomes in ADOS modules 1–3, in which lower sensitivity (SE) was found for distinctions involving children with milder ASDs (module 3 by Lord et al. 2000, SE = 0.80, versus later studies: de Bildt et al. 2009, SE = 0.64; Gotham et al. 2008, SE = 0.49; Gotham et al. 2007, SE = 0.68). The high specificity (0.82), on the other hand, means that a positive ASD classification on the ADOS is a very strong indication for a clinician to consider diagnosing ASD. Sensitivity and specificity are tightly linked and the aim of the assessment determines which one is most important. High specificity is essential when one needs to be certain that the individuals selected actually have an ASD, for instance in autism research. High specificity can, however, lead to underinclusiveness. When the aim of the assessment is to screen for ASD, high sensitivity is crucial in order not to miss any potential case. For this purpose, lower thresholds could be considered at the expense of specificity. To prevent overinclusiveness, developmental history and current daily functioning should then be carefully reviewed. As this study included only a specific ASD group and specific control groups, further research is needed to establish the optimal cut-off points on the ADOS module 4.

This study has a number of limitations that should be taken into account when interpreting the results. First, compared to studies on the psychometric properties of modules 1–3 (de Bildt et al. 2009; Gotham et al. 2007, 2008; Oosterling et al. 2010), our study has a small sample size ($n = 93$). However, it is the first study examining module 4 in an adult population with ASD compared to specific and meaningful groups. Second, we are focused on high-functioning adult males with ASD, which means results cannot be generalized to the entire ASD population. Future studies on module 4 should comprise a larger sample, including individuals with lower levels of daily functioning, since the high-functioning character of our sample may have influenced the results. On the other hand, exactly these individuals are not always recognized during childhood. Therefore, increasing knowledge on module 4 seems most important for individuals showing milder autistic symptoms. In this light, it will also be important to include a group of high-functioning adult females, who run the risk of being undiagnosed because they might be especially good at compensating their behavior (Attwood 1999; In 't Velt-Simon Thomas and Mol 2008). Third, no standardized measures were available for the clinical diagnosis of ASD, which characterizes current practice in adult psychiatry. However, the normal clinical procedure included review of developmental history and current functioning and observation. In addition, most participants

with ASD were recruited through a specialized centre. Fourth, we did use standardized measures to diagnose schizophrenia, but not to review early developmental history in this group. Therefore, we cannot eliminate the possibility that ASD was present before the onset of schizophrenia. However, this possibility is minimized by the fact that these individuals were extensively tested in a specialized psychosis centre and selected for this study by experienced clinicians. The control groups in the current sample were comparatively homogeneous and aimed to challenge the ADOS by comparing ASDs with other psychiatric groups with social deficits. For the investigation of ADOS' value in differential diagnostics, examining different subtypes of schizophrenia and other diagnostic groups will be of great relevance as well (e.g. anxiety disorder, depression, ADHD, and OCD).

In summary, the ADOS module 4 is a reliable instrument that has good predictive value for ASD. It can adequately discriminate ASD from psychopathy and typical development in an adult population. With respect to schizophrenia, discrimination is more difficult due to behavioral overlap. These groups are, however, different on some core items. Although ADOS module 4 fails to classify ASD in a significant proportion of our higher functioning and more mildly affected ASD group, its ASD classification is a strong lead for a clinician to at least consider an ASD diagnosis. Explorative analyses of the revised algorithm indicate that a revision -in line with modules 1-3 and developments in criteria for ASD- could be beneficial for discriminating ASD from schizophrenia.

Acknowledgments Thanks to all the participants for their time and effort. For assistance in recruitment and/or IQ testing: J. Ferwerda, M. Meijer, and V. Baas. Audiovisual service of Lentis for video-recordings. E. Haitel for manual support. The work was supported by a Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) VIDI grant and a Marie Curie Excellence Grant of the European Commission to CK. A version of the paper was presented as a poster at the 9th International Meeting for Autism Research (IMFAR), Philadelphia, US.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- Attwood, T. (1999). The pattern of abilities and development of girls with Asperger's syndrome. www.faaas.org/articles4.html.
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high

- functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34(2), 163–175.
- Bartels, A. A. J., & Bruinsma, C. L. (2008). Autisme spectrum stoornissen in de forensische psychiatrie. In E. H. Horwitz, C. E. J. Ketelaars, & A. M. D. N. van Lammeren (Eds.), *Autisme spectrum stoornissen bij normaal begaafde volwassenen*. Assen: Koninklijke van Gorcum.
- Bleuler, E. (1911). Dementia praecox oder gruppe der schizofrenien [Dementia praecox or the group of schizophrenias]. In *Handbuch der Psychiatrie*. Leipzig/Wien: Deuticke.
- Brugha, T., McManus, S., Meltzer, H., Smith, J., Scott, F. J., Purdon, S., et al. (2009). *Autism spectrum disorders in adults living in households throughout England: Report from the adult psychiatric morbidity survey 2007*. Leeds: The NHS Information Centre for Health and Social Care.
- Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry*, 45(4), 719–726.
- de Bildt, A., Sytema, S., Ketelaars, C., Kraijer, D., Mulder, E., Volkmar, F., et al. (2004). Interrelationship between Autism Diagnostic Observation Schedule-Generic (ADOS-G), Autism Diagnostic Interview-Revised (ADI-R), and the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) classification in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, 34(2), 129–137.
- Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *Journal of Clinical Psychiatry*, 66, 3–8.
- Frith, U., & Happé, F. (2005). Autism spectrum disorder. *Current Biology*, 15(19), 786–790.
- Giel, R., & Nienhuis, F. (1996). *Vragenschema's voor Klinische Beoordeling in Neuropsychiatrie [SCAN-2.1: Schedules for clinical assessment in neuropsychiatry]*. Groningen: World Health Organization.
- Gillberg, C. L. (1992). The Emanuel miller memorial lecture 1991. Autism and autistic-like conditions: subclasses among disorders of empathy. *Journal of Child Psychology and Psychiatry*, 33(5), 813–842.
- Goldstein, G., Minshew, N. J., Allen, D. N., & Seaton, B. E. (2002). High-functioning autism and schizophrenia: a comparison of an early and late onset neurodevelopmental disorder. *Archives of Clinical Neuropsychology*, 17(5), 461–475.
- Gotham, K., Risi, S., Dawson, G., Tager-Flusberg, H., Joseph, R., Carter, A., et al. (2008). A replication of the Autism Diagnostic Observation Schedule (ADOS) revised algorithms. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(6), 642–651.
- Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, 37(4), 613–627.
- Gray, K. M., Tonge, B. J., & Sweeney, D. J. (2008). Using the Autism diagnostic interview-revised and the Autism diagnostic observation schedule with young children with developmental delay: Evaluating diagnostic validity. *Journal of Autism and Developmental Disorders*, 38(4), 657–667.
- Hare, R. D. (1991). *The Hare psychopathy checklist-revised*. Toronto: Multi-Health Systems.
- Howlin, P. (2000). Outcome in adult life for more able individuals with autism or Asperger syndrome. *Autism*, 4(1), 63.
- In 't Velt-Simon Thomas, J. M., & Mol, A. J. J. (2008). Normaal tot hoogbegaafde vrouwen met een autismespectrumstoornis. *Wetenschappelijk Tijdschrift Autisme*, 7(3), 104–112.
- Johnson, S. A., Filliter, J. H., & Murphy, R. R. (2009). Discrepancies between self- and parent-perceptions of autistic traits and empathy in high functioning children and adolescents on the autism spectrum. *Journal of Autism and Developmental Disorders*, 39(12), 1706–1714.
- Kan, C. C., Buitelaar, J. K., & van der Gaag, R. J. (2008). Autismspectrumstoornissen bij volwassenen [Autism spectrum disorders in adults]. *Nederlands Tijdschrift Geneeskunde*, 152(24), 1365–1369.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous child*, 2, 217–250.
- Kay, S. R., Fiszbein, A., & Opfer, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, 13(2), 261–276.
- Kohn, Y., Fahum, T., Ratzoni, G., & Apter, A. (1998). Aggression and sexual offense in Asperger's syndrome. *The Israel Journal of Psychiatry and Related Sciences*, 35(4), 293–299.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One*, 2(9), e883.
- Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry*, 63(6), 694–701.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223.
- Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism diagnostic observation schedule. Manual*. Los Angeles, CA: Western Psychological Services.
- Luteijn, F., & Barelds, D. (2004). *GIT 2: Groninger intelligentie test 2*. Amsterdam: Harcourt Test Publishers.
- Noterdaeme, M., Sitter, S., Mildemberger, K., & Amorosa, H. (2000). Diagnostic assessment of communicative and interactive behaviours in children with autism and receptive language disorder. *European Child and Adolescent Psychiatry*, 9(4), 295–300.
- Oosterling, I., Roos, S., de Bildt, A., Rommelse, N., de Jonge, M., Visser, J., et al. (2010). Improved diagnostic validity of the ADOS revised algorithms: A replication study in an independent sample. *Journal of Autism and Developmental Disorders*, 40, 689–703.
- Papanikolaou, K., Paliokosta, E., Houliaras, G., Vgenopoulou, S., Giouroukou, E., Pehlivanidis, A., et al. (2009). Using the Autism diagnostic interview-revised and the Autism diagnostic observation schedule-generic for the diagnosis of Autism spectrum disorders in a Greek sample with a wide range of intellectual abilities. *Journal of Autism and Developmental Disorders*, 39(3), 414–420.
- Reaven, J. A., Hepburn, S. L., & Ross, R. G. (2008). Use of the ADOS and ADI-R in children with psychosis: Importance of clinical judgment. *Clinical Child Psychology and Psychiatry*, 13(1), 81–94.
- Robertson, J., Tanguay, P., L'Ecuyer, S., Sims, A., & Waltrip, C. (1999). Domains of social communication handicap in autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(6), 738.
- Rogers, J., Viding, E., Blair, R. J., Frith, U., & Happe, F. (2006). Autism spectrum disorder and psychopathy: Shared cognitive underpinnings or double hit? *Psychological Medicine*, 36(12), 1789–1798.
- Seltzer, M. M., Krauss, M. W., Shattuck, P. T., Orsmond, G., Swe, A., & Lord, C. (2003). The symptoms of autism spectrum disorders in adolescence and adulthood. *Journal of Autism and Developmental Disorders*, 33(6), 565–581.
- Seltzer, M. M., Shattuck, P., Abbeduto, L., & Greenberg, J. S. (2004). Trajectory of development in adolescents and adults with autism.

- Mental Retardation and Developmental Disabilities Research Reviews*, 10(4), 234–247.
- Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., et al. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 37(9), 1735–1747.
- Sheitman, B. B., Kraus, J. E., Bodfish, J. W., & Carmel, H. (2004). Are the negative symptoms of schizophrenia consistent with an autistic spectrum illness? *Schizophrenia Research*, 69(1), 119–120.
- van Lang, N. D., Boomsma, A., Sytema, S., de Bildt, A. A., Kraijer, D. W., Ketelaars, C., et al. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry*, 47(1), 37–44.
- Volden, J., & Lord, C. (1991). Neologisms and idiosyncratic language in autistic speakers. *Journal of Autism and Developmental Disorders*, 21(2), 109–130.
- Volkmar, F. R., & Cohen, D. J. (1991). Comorbid association of autism and schizophrenia. *American Journal of Psychiatry*, 148(12), 1705–1707.
- Wechsler, D. (1997). *Wechsler adult intelligence scale-III (WAIS-III)*. New York: Psychological Corporation.
- Wing, L., & Potter, D. (2002). The epidemiology of autistic spectrum disorders: Is the prevalence rising? *Mental Retardation and Developmental Disabilities Research Reviews*, 8(3), 151–161.
- World Health Organization. (1993). *International classification of diseases (ICD-10)*. Geneva: World Health Organization.